Do Psychological Fallacies Influence Trading in Financial Markets? Evidence from the Foreign Exchange Market

Michael Bleaney, Spiros Bougheas, Zhiyong Li

November, 2015

Abstract

Research in both economics and psychology suggests that, when agents predict the next value of a random series, they frequently exhibit two types of biases, which are called the gambler's fallacy (GF) and the hot hand fallacy (HHF). The gambler's fallacy is to expect a negative correlation in a process which is in fact random. The hot hands fallacy is more or less the opposite of this – to believe that another heads is more likely after a run of heads. The evidence for these fallacies comes largely from situations where they are not punished (lotteries, casinos and laboratory experiments with random returns). In many real-world situations, such as in financial markets, succumbing to fallacies is costly, which gives an incentive to overcome them. The present study is based on high-frequency data from a market-maker in the foreign exchange market. Trading behaviour is only partly explained by the rational exploitation of past patterns in the data. There is also evidence of the gambler's fallacy: a tendency to sell the dollar after it has risen persistently or strongly.

Keywords: Gambler's Fallacy, Hot hand Fallacy, Foreign Exchange Market

JEL: G02, G15

Michael Bleaney: School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom; e-mail: leamb@exmail.nottingham.ac.uk ; tel.no. +44-115-9515464

Spiros Bougheas (Corresponding Author): School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom; e-mal: spiros.bougheas@nottingham.ac.uk ; tel.no. +44-115-8466108

Zhiyong Li: Nottingham University Business School China, University of Nottingham, 199 Taikang East Road, Ningbo 315100, China; e-mail: Zhiyong.Li@nottingham.edu.cn ;tel,no. +86-574-88186465

1. Introduction

In the latter part of the twentieth century the efficient markets hypothesis came under challenge from the new field of behavioural finance, which drew on insights from psychology to analyse financial markets. Behavioural finance covers a wide range of theories (see Barberis andThaler, 2002, or Ritter (2003) for a survey, and Fama (1998) for a critical view; and Shiller (2000) or Shleifer (2000) for the limitations of speculators in preventing bubbles in stock markets). Contributions to the field in this journal include Peters (2003), Andersen (2010), Talpsepp (2011) and Das (2012). In the present paper we investigate whether psychological fallacies that have been observed in gambling environments are also in evidence in financial markets.

Research in both economics and psychology suggests that, when agents predict the next value of a random series, they frequently exhibit one of two types of bias, called respectively the gambler's fallacy (GF) and the hot hand fallacy (HHF) (for a survey, see Ayton and Fischer, 2004). The gambler's fallacy is to expect a negative correlation in a process which is in fact random (for example, to believe that the next toss of a fair coin is more likely to be tails after a run of heads). The hot hand fallacy, which works in the opposite direction, is to expect that another toss of heads is more likely after a run of heads. These fallacies may seem mutually contradictory, but the empirical evidence suggests that they apply in slightly different contexts, with the HHF more likely to occur where human performance is thought to be involved. For example, Croson and Sundali (2005) find that, amongst gamblers at a roulette wheel in a casino, a streak of one colour is more likely to induce bets on the opposite colour (GF), but that bets are made on more numbers after winning than losing, as if the gambler's own performance is positively serially correlated (HHF). Though there is plenty of evidence both from laboratory experiments and from the field (lotteries and casinos) on the GF and the HFF, most of it is limited to purely random environments with no variation in signal strength (e.g. the toss of a coin). Moreover the probability of winning is not adversely affected by players' psychological biases, so in these environments we cannot test whether players learn to overcome these biases when there is an incentive to do so.

Financial markets are especially interesting because the wrong decisions are potentially very costly, so they offer a strong incentive for traders not to succumb to psychological biases in decision-making. Unlike at the roulette wheel or in the choice of lottery numbers, where players' choices make no difference to the chance of success because the underlying process is random, in financial markets such biases can prove to be very costly if they do not match the underlying dynamics of prices. If people learn to overcome psychological biases in environments where such biases are punished, these biases should not be observed in financial markets. Moreover financial markets offer a particularly rich field of investigation in that the behaviour of participants might be affected by characteristics of the data other than the direction of the signal (whether the price is moving up or down), such as, for example, the signal strength (how fast the price moves). On the other hand, financial markets present an extra challenge because the underlying data generating process is unknown, and the research design needs to take account of the fact that some behaviour may represent the rational exploitation of patterns in the past data.

Up to now researchers have studied the behaviour of investors by comparing market outcomes with investor beliefs elicited through surveys.[1] Asset price and returns in these studies reflect market outcomes and cannot be pinned down to specific trades. In this paper, we take an alternative approach and use high-frequency data from the foreign exchange market (the same data as used in Evans and Lyons, 2002). The advantage of this approach is that we have data on specific trades and on the movements of the exchange rate in the period leading up to each trade. Although our data set may seem somewhat old, since it dates from 1996, an important feature of it is that there was no algorithmic trading at the time, so we can be confident that all the trades were executed by human beings rather than computer programmes.

The dataset gives the time, direction of trade (buy or sell the US dollar) and the price of the trade. We interpret the psychological fallacies discussed above as hypotheses about the effect of exchange rate movements up to trade T on the direction of trade T (we do not have data on the size of the transaction). We define a streak as a sequence of movements of the estimated mid-price of the dollar in the same direction over successive transactions, and we investigate how the characteristics of a streak affect the buy/sell decision. The characteristics of the streak are its length (the number of transactions for which the streak has lasted) and its width (the average movement in price per transaction over the streak). The length of the streak is a measure of the persistence of exchange rate movements, and the width is a measure of the speed of these movements. If the behaviour of traders in the market is influenced by GF and/or HHF, we expect the orders placed to be affected by the properties of the streak. For example, the gambler's

[1] For the advantages of survey data see Manski (2004). For studies using data from capital markets see Dominitz and Manski (2007, 2011) and Amromin and Sharpe (2014) and for a study that uses data from the foreign exchange market see Jongen *et al.* (2012).

fallacy would suggest that persistent upward movement of the dollar would make dollar sales more likely (which we term "trading against the trend"). In other words, we expect the buy/sell decision to be predictable to some extent from the streak characteristics.

Our analysis does not rely on the assumption that the trades in the dataset are observed by all market participants (they are not). Traders can easily obtain high-frequency information about how exchange rates are evolving from dealers' quotes and computer screens, so the exchange rate information in our dataset is effectively public. Thus in our view the dataset may be interpreted as traders' reactions to high-frequency price movements.

Our results show that there is a significant tendency to trade against the streak (i.e. to bet that the streak will end), and that this effect increases with the length and width of the streak. This is apparently consistent with traders suffering from the gambler's fallacy. However, it is possible that streaks in exchange rates are indeed followed by movements in the opposite direction, and that traders are simply exhibiting rational behaviour, trying to exploit a pattern that they have observed in past data. We implement various tests to investigate whether this rational explanation of our results stands up to scrutiny. Our results are mixed. It is not clear that streak variables explain future exchange rate movements, as the rational explanation requires. On the other hand, we find that the predictions from an autoregressive model of exchange rate returns help to explain trading behaviour. Nevertheless the streak variables retain their statistical significance, which suggests that psychological factors also play a role.

The rest of the paper is organised as follows. In Section 2, we review related studies from both the economic and the psychology literatures. In section 3, we describe the raw data and formally define the streak length and the streak width. In Section 4, we analyse the relationship among traders' behaviour (the order flow), streak length and width using a probit model. Our results in Sections 3 and 4 uncover a statistically significant relationship between the trend and the behaviour of traders. In Section 5, we control for irregular time spaces and past order flows to examine the robustness of our results. In Section 6, we examine whether the behaviour is consistent with rationality. We conclude in Section 7.

2. Related Literature

Both the gambler's fallacy and the hot hand fallacy are cognitive biases arising when individuals predict the future outcomes of a random series.

Agents who suffer from GF expect a positive (negative) shock after a negative (positive) shock. Clotfelter and Cook (1993) find that lottery players exhibit GF in a numbers game: numbers which have already been drawn become very quickly less popular among the players.

In contrast, agents who suffer from HHF expect a positive (negative) shock after a positive (negative) shock. Gilovich et al. (1985) have demonstrated the presence of HHF among basketball players and fans, who believe (erroneously) that a player who is having a streak of successes also has a high probability of scoring the next time. Camerer (1989) and Brown and Sauer (1993) also find evidence on HHF in the basketball betting market. Guryan and Kearney (2008) provide evidence for the "lucky store effect", which refers to the tendency of gamblers to believe that there is a higher probability of winning if they buy a ticket from a store which has recently sold a winning ticket.

Though GF and HHF make contradictory predictions, researchers find that both can be observed in the behaviour of the same individual. Sundali and Croson (2006) and Croson and Sundali (2005) identify both effects in the behaviour of the same single individual by using data collected from a casino. Individuals tend to shun recently winning outcomes (GF), but bet more frequently after successful bets (HHF). Psychologists and economists offer several explanations for this phenomenon.

2.1. Psychological Theories

The psychology literature provides two explanations for the relationship between GF and HHF. One explanation suggests that both fallacies are cognitive biases caused by lack of knowledge of probability theory. The other explanation distinguishes two different types of expectations with respect to human performance and natural events (Ayton and Fischer, 2004). In human performance, individuals normally exhibit positive recency (HHF). In natural events, individual normally exhibit negative recency (GF).

Tversky (1974) suggests that GF, one of the representativeness heuristic biases, is caused by poor understanding of probability theory. Individuals tend to expect a small sample to have the statistical characteristics of a large sample. Consequently, when individuals toss a fair coin they attach a probability of more than 50% of getting a tail after observing a streak of heads. On the other hand, when the streak of heads becomes very long or when the frequency of heads is consistently larger than that of tails, individuals may reject the belief that the series is random and start to believe that the outcome of any toss is positively related to the previous one. Then GF becomes HHF. Rao (2009) has designed an experiment to study the relation between GF and HHF. The experiment suggests that GF normally happens after short streaks and HHF happens after long streaks. The experiment also identifies the transition from GF to HHF. Guryan and Kearney (2008) argue that the "lucky store" effect challenges this interpretation, but in fact it seems consistent with HHF, since stores with long losing streaks are avoided.

Ayton and Fischer (2004) give an alternative explanation. They argue that the representativeness explanation cannot show why HHF is observed in basketball games but the GF is not, and why the GF is observed in lottery games but HHF is not. In their experiments individuals suffer from GF when predicting a series of natural events such as the toss of a coin or the spin of a roulette wheel, and suffer from HHF when predicting a series of human activities such as basketball players' shooting success or whether the first serve of a tennis player is returned.

2.2. Economic Explanations

Economists are interested in GF and HHF because the two biases may help to explain puzzles or anomalies in financial markets. The model of Rabin and Vayanos (2010) (hereafter RV) follows the suggestion of the psychological literature and tries to link the two fallacies. The RV model assumes that agents believe that random shocks are negatively serially correlated (in other words, they suffer from GF). GF becomes HHF after a long streak because agents update their beliefs about the state of the world in an erroneous manner.

In the RV model, agents can observe a signal (x) which is generated by a mean-reverting process and an i.i.d shock: $x_t = \alpha x_{t-1} + u_t$. The parameter α is unknown to the agents. Agents erroneously believe that u_t is negatively autocorrelated, which generates a bias that is akin to GF. As agents learn from observing outcomes, it is assumed that they update their beliefs about α, but also (crucially) that they fail to correct their erroneous belief about the shocks. After a succession of positively correlated shocks, agents may revise their beliefs about α upwards so much that they believe that α is greater than one. When this happens their expectations become extrapolative, as if agents suffered from the HHF. In effect agents correct for their erroneous beliefs about shocks by overestimating α, and the prediction process is a combination of a GF bias and this compensating HHF bias. The model predicts GF behaviour for short streaks and HHF behaviour for long streaks.

Barberis et al. (1998) (hereafter BSV) assume that agents believe that random movements of an asset price have two regimes: reversing (which can be interpreted as GF) and trending (which can be interpreted as HHF). Agents continuously update their subjective probabilities (relative strength) of the two regimes. GF becomes HHF after a long streak, as in the RV model, because the relative strength of GF and HHF changes as the streak increases. As in the case of the RV model, traders have erroneous beliefs that they cannot entirely correct through learning. Conservatism means that individuals update their beliefs slowly when facing new information. Traders believe that the series, which is actually random, follows either one of two regimes with given parameters: mean-reverting or trending (continuation in the same direction). The prediction is a weighted average of the predictions of the two regimes, with the weights being adjusted in an optimal fashion in response to past data. Thus, as in the RV model, agents adjust their beliefs in the light of experience, but never arrive at the correct data-generating process. The model suggests that traders under-react to short term shocks and over-react to long term trends. The mean-reverting regime can be considered as the GF effect and the trending regime can be explained as the HHF effect.

Both models can be criticised theoretically on the grounds that what agents can or cannot learn is somewhat arbitrary, but they are supported by experimental evidence when compared to either the random walk model or the Bayesian learning model (Bloomfield and Hales, 2002; Asparouhova et al., 2009). However, overall, the RV model does a better job than the BSV model.

3. Data and Series Construction

3.1. Data

In this section, we describe the dataset and the methodology we employed for generating the series which we used in our analysis.

The data include all the interbank tick-by-tick prices and order flows of the Deutsche Mark (DEM) and the Japanese yen against the US dollar from May 1 to August 31 1996 on the Reuters Dealing 2000-1 (hereafter D2000-1). We report results for the DEM only; our findings for the Japanese yen were similar. The data set format is shown in Table 1. Trades on D2000-1 happen between two anonymous dealers: a calling dealer, who requires quotes, and a quoting dealer. The quoting dealer offers bid and ask prices to the calling dealer. The calling dealer has to make a quick decision to buy dollars (make a positive order flow) or sell dollars (make a negative order flow) or reject the quote. If a transaction is made, the time and the direction will be recorded by the system. Two things need to be mentioned. First, traders can only observe their own trading records. Second, though both bid and ask prices (two series of exchange rates) were quoted by the calling dealers, only the price that reflects the direction of actual trade is in the dataset (and the price may be slightly more favourable to the trader than the quote).

Although the Reuters D2000-1 system does not publish transaction prices, this does not affect our analysis. Traders in the market are aware of prices at all times for several reasons. First, traders in the market frequently ask for quotes to obtain the latest prices and to seek arbitrage opportunities. The no-arbitrage condition, on which the Evans and Lyons (2002) model relies, requires that quotes of different market makers are identical at a given time. Second, users of the Reuters D2000-1 system have indicative FXFX data (real time indicative prices) for reference. Although FXFX data are not transactions data, they do reflect the basic tendencies of prices. Third, users of the Reuters D2000-2 system have data from Reuters D2000-1. Therefore, traders have many indirect methods of obtaining real-time price information, and it is reasonable to assume that transaction prices are known by everyone in the market. Moreover, if the trades recorded in our data set differ significantly from quoted prices at the time, one side of the trade is knowingly making a significant loss, which would be irrational, so one can assume that such trades are unlikely to occur.

The average time between trades is 26 seconds, with a standard deviation of 76 seconds, but the distribution is heavily skewed to the right, with a maximum of 10560 seconds (in a period when all the major financial markets are closed), and a minimum of zero. The median time between trades is nine seconds, and the interval exceeds one minute in only 9% of the sample.

3.2. Estimation of the Spread

Since our main concern is the reaction of trading decisions to the recent trend in prices, it is essential that this trend is measured accurately. The observed transaction prices are affected by the buy-sell spread, and to estimate the underlying price trend we need to correct for this effect.

Let x_t be the logarithm of the nominal exchange rate of the US dollar against the German mark for transaction t (an increase representing an appreciation of the US dollar), and let OF_t be the order flow which is equal to $+1$ when an agent buys dollars and equal to 0 when an agent sells dollars. Then

$$x_t = \begin{cases} A_t, \ Buy\ order\ (OF_t = 1) \\ B_t, \ Sell\ order\ (OF_t = 0) \end{cases} \tag{1}$$

where B_t denotes the logarithm of the bid price and A_t denotes the logarithm of the ask price and $A_t \geq B_t$. The unobserved bid-ask spread SP_t is the difference between the ask and the bid prices,

$$SP_t = A_t - B_t.$$

The spread distorts the trend in prices: if a buy order is followed by a sell order, or *vice versa*, the change in *x* will not reflect the true price trend. We therefore need to estimate and eliminate this effect.

We define the mid-price for transaction t, M_t, as the average of the bid and the ask prices. More formally,

$$M_t = \frac{1}{2}(A_t + B_t) = \begin{cases} A_t - \frac{SP_t}{2}, (OF_t = 1) \\ B_t + \frac{SP_t}{2}, (OF_t = 0) \end{cases}. \tag{2}$$

Under the supposition that the bid-ask spread is fixed, the change in the mid-price is given by:

[Please insert Table 1 about here]

$$\Delta M_t = \frac{1}{2}[(A_t - A_{t-1}) + (B_t - B_{t-1})] = A_t - A_{t-1} = B_t - B_{t-1}. \tag{3}$$

where Δ denotes the first-difference operator. Since in this case the spread is unknown, it has to be estimated from the data, which we do by using the model of Huang and Stoll (1997).[2] Using the estimate of the spread ($\widetilde{SP}$) we can adjust the observed changes in the transactions price returns using the following formula:

$$\Delta s_t = \Delta x_t - \widetilde{SP}_t(OF_t - OF_{t-1}), \tag{4}$$

where *s* is the estimate of the mid-price.

[2]This is a standard model when the direction of trade is known, as here. For details of the estimation see Bleaney and Li (2014).

3.3. Streak Length and Width

If agents believe that a given series is autocorrelated, then the properties of past observations should have predictive power over agents' actions (in this case to buy or sell the US dollar). In our analysis we will focus on two properties of a streak, which is defined as a sequence of movements of the estimated mid-price of the dollar in the same direction over successive transactions.

The first property is the streak length, k_t, which denotes the number of transactions up to transaction *t* that the series continues to move in the same direction. The streak length is a non-negative integer.

We also need a measure of the strength of the streak, and a candidate would be the distance between the current exchange rate and the one at the beginning of the streak. However, clearly the streak length and the distance are positive correlated. In order to avoid this problem we use the streak width, w_t, defined as the streak distance divided by the streak length.

We can formally express the streak length as:

$$k_t = \begin{cases} k_{t-1}, & sign[\Delta s_t] = sgin[\Delta s_{t-1}] \\ 0, & sign[\Delta s_t] \neq sign[\Delta s_{t-1}] \end{cases}, \quad (5)$$

where

$$sign_t = \begin{cases} 1, & \Delta s_t > 0 \\ sign_{t-1}, \; if & \Delta s_t = 0 \\ -1 & \Delta s_t < 0 \end{cases}$$

The formal expression for the streak width is given by:

$$w_t = \frac{100|s_t - s_{t-k_t-1}|}{k_t + 1}. \quad (6)$$

If traders suffer from GF and/or HHF, their choices should be influenced by the streak length and the streak width. For example, if the dollar has been rising for some time (a long streak) or particularly fast (a wide streak), the HHF would predict that the next transaction will tend to be a dollar purchase. Most of the evidence relating to these fallacies has been gathered from coin-tossing experiments where, by definition, the streak width is constant. Consequently, the literature has focused on the effects of streak length on the behaviour of agents and has ignored streak width. That is fine when the evidence is related to scoring in a basketball game or gambling in a numbers game, but is not necessarily appropriate for the foreign exchange market where the shocks vary in size.

Table 2 presents the summary statistics of the data. The streak length (k) and streak width (w) are calculated from the series of estimated mid-price exchange rate returns (Δs).

[Please insert Table 2 about here]

Figure 1 shows the histogram of streak length. The average streak length is less than one, and the fraction of instances where $k_t = 0$ is greater than 50%, which suggests that the sign of exchange rate returns changes very frequently. Figure 3 shows the histogram of streak width, which is also strongly skewed to the right.

[Please insert Figure 1 about here]

4. Streak Characteristics and Trading Behaviour

In this section we report some empirical tests of whether the characteristics of a streak are correlated with trading decisions in a way that is consistent with the literature on psychological fallacies. Specifically, we test whether streak length and width are correlated with order flows as the GF or HHF would predict.

We wish to test whether trading behaviour is consistent with the GF, with the HHF or with neither. To do this we use a modified indicator of trade direction:

$$TF_t = \begin{cases} 1 & if \ \frac{2(OF_t - 1/2)}{sign_t} = 1 \\ 0 & \frac{2(OF_t - 1/2)}{sign_t} = -1 \end{cases},$$

When $TF_t = 1$ the trader trades with the trend (buys the \$ when it has just risen or sells the \$ when it has just fallen) and when $TF_t = 0$ the trader trades against the trend. Formally, $TF = 1$ if $OF = 1$ and $sign = +1$ or if $OF = 0$ and $sign = -1$. Conversely, $TF = 0$ if $OF = 1$ and $sign = -1$ or if $OF = 0$ and $sign = +1$. The HHF would tend to predict $TF = 1$, particularly in the case of a long streak, whereas the GF would predict $TF = 0$.

We can use the following regression to study the impact of the streak length on trend-following behaviour:

$$TF_t = \Phi(\beta_1 + \beta_2 k_t) + \varepsilon_t. \tag{7}$$

[Please insert Figure 2 about here]

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. If the coefficient β_2 is positive, then the probability of trend-following behaviour increases as the streak length becomes longer (as the HHF would predict) and vice versa. A negative β_2 would be consistent with the GF.

The first column (7) of Table 3 reports the results, while the second column (7*) reports the average marginal effects. The coefficient β_2 is significant and negative, which indicates that trend-following becomes less probable as the streak length increases. The marginal effects suggest that for each increase of one in the streak length, the probability of following the trend on average decreases by 13.7%. These results are consistent with the hypothesis that trading decisions reflect the GF.

We can examine the relationship between streak length and trading decisions more closely by replacing streak length by a series of dummies (Dk) for different streak lengths such that

$$Dk_i = \begin{cases} 1 & if \quad k_t = i \\ 0 & \quad\;\; k_t \neq i \end{cases}.$$

The probit model is written as:

$$TF_t = \Phi(\beta_1 + \sum_{i=1}^{n} \delta_i \, Dk_{it}) + \varepsilon_t. \tag{8}$$

Compared with (7), (8) uses dummies for the streak length rather than the streak length itself, which allows the influence of the streak length on order flows to be non-linear. This non-linear relationship will be captured by the coefficients of the dummies. Since the proportion of observations with streak length over five is rather small, we amalgamate all the dummies for a streak length of five or more into one dummy for a streak length of five or more (Dk_{5+}).

The third and fourth columns of Table 3 report the results of estimating equation (8) and the marginal effects (denoted by 8*) respectively. The results are similar to those for equation (7): as the streak length increases, the probability of following the trend decreases, and this occurs at each step, which indicates that equation (7) is a good approximation to the underlying pattern. For example, from (8*), compared with the case when the streak length is zero, the probability of following the trend is 23.3% lower when the streak length is one and 29.0% lower when the streak length is two. Furthermore, the decline in the probability of trend-following falls as the streak length increases: when the streak length increases from one to two, the probability of trend-following falls by 5.7%, and when the streak length increase from two to three, the probability of trend-following falls by 2.1%. Figure 3 plots the predicted probability of trend following against the streak length implied by equation (8) in Table 3.

[Please insert Figure 3 about here]

Trend-following behaviour when the streak length is very short (less than zero or equal to one) might arise from the sequential nature of order arrivals. Traders holding the same expectations with regard to exchange rate movements may place orders in the same direction almost simultaneously, which causes two adjacent orders to be positively correlated. For example, suppose that the streak length is zero and the exchange rate rises in period t, and that both traders, A and B, believe that the exchange rate will rise in the future. The buy order of trader A comes first. The buy order of trader B comes after A's order, when (we shall assume)

[Please insert Table 3 about here]

the exchange rate has started to rise. As a result, B's order is recorded as a trend-following order with a streak length of one. Thus, almost simultaneous actions by traders reacting to similar information might be responsible for some apparent trend-following behaviour. This effect will tend to work against the GF.

4.1. Streak Width

In this section, we investigate whether the strength as well as the direction of the signal appears to make a difference to trading decisions. We do this by adding the streak width to equation (7):

$$TF_t = \Phi\left(\beta_1 + \beta_2 w_t + \sum_{i=1}^{5} \delta_i \, Dk_{it}\right) + \varepsilon_t, \quad (9)$$

where β_2 measures the effect of the streak width on trend-following behaviour.

The results and the marginal effects are reported in Table 4.[3] The streak-width coefficient is negative and significantly different from zero, which suggests that the streak width has an additional negative effect on trend-following behaviour, for a given streak length. On the other hand its *Z*-statistic is quite small, considering the sample size, and the estimated effect of a one-standard-deviation increase in streak width on the probability of trend following is only -1.4% (= -0.479 x 0.0292). Thus streak width is of minor importance compared to streak length.

[Please insert Table 3 about here]

5. Robustness checks

5.1. Time intervals

One of the important properties of the data is that time is measured in trades, which means that the time interval between consecutive observations is not identical. To some extent this is an advantage, because periods when there is little trading, which may be atypical, are represented by few observations and should therefore have less influence on the results. Nevertheless it is important to check that our results are not sensitive to the time interval between trades.

Let T_t be the number of seconds between orders *t* and *t-1*. For example, the first trade in Table 1 occurred at 18:45:40, and the second trade occurred at 18:46:23. There are thus 43 seconds between the two trades, and accordingly $T_2 = 43$. As a first robustness test, we allow all the coefficients in equation (9) to vary with *T*. Thus we estimate the following probit regression:

$$TF_t = \Phi\left(\beta_1 + \beta_2 w_t + \sum_{i=1}^{5} \delta_i Dk_{it} + \beta_3 T_t + \sum_{i=1}^{5} \gamma_i T_i Dk_{it}\right) + \varepsilon_t. \qquad (10)$$

As a second robustness test, we exclude the data with the shortest and the longest intervals between trades. To be specific, we re-estimate equation (9) using only the trades between the 25th and the 75th percentiles for *T* (respectively three seconds and 23 seconds). Table 5 reports the results of these two tests and the corresponding marginal effects, respectively.

[3] An anonymous referee has suggested that we should also take into account brief interruptions to streaks, which would otherwise have been significantly longer. To address this point, we have also estimated the model with the exchange rate series smoothed by moving averages of up to five periods, which should eliminate such interruptions. The results are very similar to those reported in Table 4.

In the first test, (columns 10a and 10a*) the dummy variable interacted with the time interval is significant at the 5% level in only one out of the five cases (k = 1). Thus the estimated coefficients do not seem to vary much with the time interval between trades. In the second case (the truncated sample – columns 10b and 10b*), the estimated marginal effects are very similar to those shown in Table 4.

Lastly, it is possible that frequent trades might not have been recorded in the exact order. Wrong ordering of trades can have a marked effect on the streak variables. We have confirmed that regressions using only periods of low trading volume, when such mis-ordering should be less likely, yield similar results as those for the whole sample. (The results are presented in Appendix A.1.)

5.2. Microstructure

According to Evans and Lyons (2002), there is "hot-potato" trading on the Reuters D2000-1 system; traders pass their inventory imbalance onto others, and thus current trading behaviour is influenced by past trading behaviour. To allow for this, we add lagged trade direction to the model:

$$TF_t = \Phi\left(\beta_1 + \beta_2 w_t + \sum_{i=1}^{5} \delta_i \, Dk_{it} + \beta_4 TF_{t-1}\right) + \varepsilon_t \tag{11}$$

Table 6 reports the results and the corresponding marginal effects. The coefficient of TF_{t-1} is significant and positive, as the hypothesis of hot-potato trading would predict. More important for our purposes is the fact that including a lag of the dependent variable produces estimates of the other coefficients that are similar to those shown in Table 4, so that our previous results seem to be robust to controlling for hot-potato trading. The results are also robust when we control for more lags for trade direction (TF_{t-2}, TF_{t-3} ...). (The results are presented in Appendix A.2.)

[Please insert Table 5 about here]

[Please insert Table 6 about here]

6. Do Traders Behave Rationally?

In the last two sections, we have found strong evidence that traders tend to trade against the trend. This behaviour is consistent with the influence of the gambler's fallacy: traders observe that a currency is moving in one direction, and they tend to trade as if they were betting on that movement being reversed. This may, however, be entirely rational behaviour, and not the result of psychological fallacies at all, if this is a pattern that is observable in past data. To test this, we need to examine the dynamics of the exchange rate.

We consider two types of model of the exchange rate that traders might use. Model A is a standard autoregressive model, in which exchange rate movements are a function of up to 15 lags of themselves. Model B assumes that exchange rate movements are determined by (signed) past streak length and width, i.e. the same characteristics of the series as we used earlier in testing for psychological fallacies. The issue is whether predictions from these models explain trading behaviour better than the streak variables in Table 4. We use two types of prediction. Prediction One consists of the fitted values from Model A or B, and thus takes account of both the size and the direction of predicted exchange rate movements. Prediction Two is a binary variable based on the sign of these fitted values only. Although Prediction Two discards information about the predicted size of exchange rate movements, retaining only the predicted direction, one advantage of using Prediction Two rather than Prediction One is that it reduces the multicollinearity in the case of Model B, where the forecasting model contains variables that are closely related to those in Table 4.

A further complication is that trading tends to have a price impact, and this by itself tends to make trading look as if it is based on correct predictions of exchange rate movements, when in fact it may not be. We can control for this effect by the method of Huang and Stoll (1997) of using the residuals of a regression of exchange rate movements on order flows in place of the exchange rate movements themselves in estimating Models A and B, and we also present some results based on this alternative model. The regression is:

$$\Delta s_t = \beta_1 OF_t + \beta_2 OF_{t-1} + \beta_3 OF_{t-2} + \varepsilon_t \qquad (12)$$

In other words we replace Δs by ε in estimating Models A and B. We do not report these results because they are very similar to the results for Δs.

Table 7 shows the results of estimating Model B for short-run exchange rate movements (i.e. the movement in the price between two successive trades). In the first column of Table 7 exchange rate movements are a function of streak length only. Streak length is significant with a negative coefficient, but the R-squared is only 0.07. In the second column streak width is added, and this variable is quite a bit more significant than streak length; both variables have negative coefficients, and the R-squared has risen to 0.19. If traders are rational, the streak variables should have the same signs in the exchange rate regression as in the previous regression for trade direction (Tables 3 and 4). The rational explanation for our previous results would then be that traders bought (sold) the dollar at time *t* because they predicted from past data that the dollar was likely to rise (fall) between times *t* and *t*+1. Since the coefficients in Table 7 have the same signs as in Tables 3 and 4, rational trading is not going to be easy to distinguish from gambler's fallacy effects. The final column of Table 7 shows that slightly longer-term exchange rate movements (between trades *t*+1 and *t*+6) are not predicted by streak length or width.

[Please insert Table 7 about here]

We do not show the results for the autoregressive model of exchange rate returns (Model A), but it provides a significantly better fit, with an R-squared of 0.37 (the results are presented in Appendix A.3.), which suggests that the rational trader should prefer Model A. Unlike Model B the autoregressive model also has a significant capacity to explain longer-term returns: the R-squared for an AR(15) model of $(s_{t+6} - s_{t+1})$ is 0.34 (the results are presented in Appendix A.3.), and for an AR(100) model the R-squared rises to 0.42.

In Table 8 we add the fitted values from Models A and B (including streak width in the latter case) to the Table 4 model of trading. The predictions from each model are significant with a positive sign, suggesting that trading is at least partly rational, but the streak variables remain significant. There is an important difference between streak length and streak width, however: whilst the streak length dummies still have significant negative coefficients with marginal effects not much smaller than in Table 4, streak width now has a positive coefficient. This means that its significant negative coefficient in Table 4 can be entirely attributed to rational trading, rather than psychological fallacies. It is also notable that standard errors are much higher for Model B than for Model A, because of multicollinearity problems.

[Please insert Table 8 about here]

In Table 9, the predicted exchange rate changes from Models A and B are replaced by their sign, so that the prediction is a binomial variable (Models C and D respectively). For Model A this results in a decisive rejection of the rational model, because the sign of the prediction has a *negative* coefficient. For Model B, the multicollinearity problem evident in Table 8 is reduced, and the sign of the prediction has a significant positive coefficient. In results not shown, we have also tested the predictions of $(s_{t+6} - s_{t+1})$ from the autoregressive model, but these predictions perform very poorly, coming out with a negative coefficient.

[Please insert Table 9 about here]

In short, the streak variables continue to explain trading patterns even when we allow for rational trading based on past patterns in the data. Although streak variables help to predict future returns, they do so markedly less well than a conventional autoregressive model, so it is reasonable to assume that the autoregressive model is a more accurate representation of rational traders' behaviour.

7. Conclusion

There is evidence from previous research that people suffer from the gambler's fallacy and/or the hot hand fallacy when they predict future values of uncertain series. However, the evidence comes from laboratory and field experiments from casinos and lotteries where the underlying series is the product of a random process. Since the outcome is then entirely random, there is no cost attached to succumbing to these fallacies: the player's chances of winning remain the same. A natural question to ask is whether people learn to overcome their tendency to succumb to psychological fallacies in environments where mistakes are punished. To this end, we have studied the behaviour of professional traders in the foreign exchange market. Moreover, many studies focus only on the direction of the signal with no variation in signal strength; in financial markets, as in many other real-world situations, the strength of the signal (the speed of price movements) is potentially as important as its direction.

We have shown that the actions of traders depend on the two properties of the trend in the series of price returns, namely the streak length and the streak width, in a manner that is consistent with the gambler's fallacy (but not the hot hand fallacy). Persistent (and to a lesser extent fast) upward (downward) movements in the dollar tend to be followed by dollar sales (purchases). We then asked whether our results could be explained by traders' rational exploitation of non-random patterns in exchange rate returns, by including predictions from two models of exchange rate returns (one based on streak variables and a more conventional autoregressive model) in our model of trading. The rational models (and particularly the autoregressive model) have a significant capacity to predict future exchange rate returns. If the models that we estimate approximate to those used by rational traders, predictions from these models should help to explain trading behaviour. In our results, this is only partially true: the streak variables retain their statistical significance in every specification, but the predictions from a rational model do not always have the expected positive sign. The evidence presented here suggests that the gambler's fallacy plays a role in trading decisions in foreign exchange markets.

References

Andersen, J.V. (2010). Detecting anchoring in financial markets, Journal of Behavioral Finance 11, 129-133

Amromin, G. and S. A. Sharpe (2014). From the horse's mouth: Economic conditions and investor expectations of risk and returns. *Management Science* 60, 845-866

Asparouhova, E. N., M. G. Hertzel, and M. L. Lemmon (2009). Inference from streaks of random outcomes: Experimental evidence on beliefs in regime shifting and the law of small numbers. *Management Science* 55, 1776-1782

Ayton, P. and I. Fischer (2004). The hot hand fallacy and the gambler’s fallacy: two faces of subjective randomness? *Memory and Cognition* 32, 1369–78

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics* 49, 307–343

Barberis, N. and R. Thaler (2002). A survey of behavioral finance, NBER Working Paper no. 9222

Bleaney, M. and Z. Li (2014). Decomposing the bid-ask spread in multi-dealer markets. *University of Nottingham School of Economics Discussion Papers* 14/03

Brown, W. O. and R. D. Sauer (1993). Does the basketball market believe in the hot hand? comment. *American Economic Review* 83, 1377–1386

Camerer, C. F. (1989). Does the basketball market believe in the ‘hot hand?’ *American Economic Review* 79, 1257–1261

Clotfelter, C. T. and P. J. Cook (1993). The “gambler’s fallacy” in lottery play. *Management Science* 39, 1521–1525

Croson, R. and J. Sundali (2005). The gambler’s fallacy and the hot hand: Empirical data from casinos. *Journal of Risk and Uncertainty* 30, 195–209

Das, A. (2012). Estimating the loss from the disposition effect: a simulation study, Journal of Behavioral Finance 13, 1-10

Dominitz, J. and C. F. Manski (2007). Expected equity returns and portfolio choice: Evidence from the health and retirement study. *Journal of the European Economic Association* 5, 369-379

Dominitz, J. and C. F. Manski (2011). Measuring and interpreting expectations of equity returns. *Journal of Applied Econometrics* 26, 352-370

Evans, M. D. D. and R. K. Lyons (2002). Order flow and exchange rate dynamics. *Journal of Political Economy* 110, 170–180

Fama, E.F. (1998). Market efficiency, long-term returns, and behavioral finance, Journal of Financial Economics 49, 283-306

Gilovich, T., R. Vallone, and A. Tversky (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology* 17, 295–314

Guryan, J. and M. S. Kearney (2008). Gambling at lucky stores: Empirical evidence from state lottery sales. *American Economic Review* 98, 458–473

Huang, R. D. and H. R. Stoll (1997). The components of the bid-ask spread: A general approach. *Review of Financial Studies* 10, 995–1034

Jongen, R., W. F. C. Verschoor, C. C. Wolff, and R. C. J. Zwinkels (2012). Explaining dispersion in foreign exchange expectations: A heterogeneous agent approach. *Journal of Economic Dynamics and Control* 36, 719-735

Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72, 1329-1376

Peters, E. (2003). Simple and complex market inefficiencies: integrating efficient markets, behavioral finance, and complexity, Journal of Behavioral Finance 4, 225-233

Rabin, M. and D. Vayanos (2010). The gambler's and hot-hand fallacies: Theory and applications. *Review of Economic Studies* 77, 730–777

Rao, J. M. (2009). Inference of Subjective Sequences: When the gambler's fallacy becomes the hot hand fallacy. Yahoo! Research Labs, Mimeo

Ritter, J.R. (2003). Behavioral finance, Pacific-Basin Finance Journal 11, 429-437

Shiller , R. (2000). Irrational exuberance, Princeton, NJ: Princeton University Press

Shleifer A. (2000). Inefficient markets, Oxford: Oxford University Press

Sundali, J. and R. Croson (2006). Biases in casino betting: The hot hand and the gambler's fallacy. *Judgment and Decision Making* 1, 1–12

Talpsepp, T. (2011). Reverse disposition effect of foreign investors, Journal of Behavioral Finance 12, 183-200

Tversky, A. and D. Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases. *Science* 185, 1124-1131

Appendix

A.1. The Overlapping Issue During Peak Trading Times

The time of each trade is recorded accurately, but the short delay between the placement of an order and its execution may vary, so the data set may misrepresent the sequence in which orders were placed. This is obviously more of an issue when trading is more frequent.

To overcome this, we use a subset of the data where the time between trades t and t-1 and also between trades t-1 and t-2 is greater than 30 seconds. In this case the sequence of orders should definitely be accurate. Since the results (reported below) are similar to those for the whole data set, we conclude that our results are not distorted by variation in the time elapsed between placement and execution of orders. An interesting thing is that we observe the coefficient of DK_5 being smaller than the DK_4. Table A1 shows the results for the following regressions:

$$TF_t = \Phi\left(\beta_1 + \sum_{i=1}^{n} \delta_i\, Dk_{it}\right) + \varepsilon_t \quad when\ T_{t-1} > 30\ and\ T_t > 30 \quad (A1)$$

$$TF_t = \Phi\left(\beta_1 + \sum_{i=1}^{n} \delta_i\, Dk_{it}\right) + \varepsilon_t \quad when\ T_t > 30\ (A2)$$

$$TF_t = \Phi\left(\beta_1 + \beta_2 w_t \ + \sum_{i=1}^{5} \delta_i\, Dk_{it}\right) + \varepsilon_t \quad when\ T_{t-1} > 30\ and\ T_t > 30\ (A3)$$

$$TF_t = \Phi\left(\beta_1 + \beta_2 w_t \ + \sum_{i=1}^{5} \delta_i\, Dk_{it}\right) + \varepsilon_t \quad when\ T_t > 30\ (A4)$$

[Please insert Table A1 about here]

A.2. Market Microstructure

Further to the regression in Section 5.2, we control for more lags for trade direction (TF_{t-2}, TF_{t-3} ...). The results, which are shown in Table A2, are very similar to those in Table 6, which suggests that our results are robust.

$$TF_t = \Phi\left(\beta_1 + \beta_2 w_t \ + \sum_{i=1}^{5} \delta_i\, Dk_{it} + \sum_{j=1}^{10} \beta_{3+j} TF_{t-i}\right) + \varepsilon_t$$

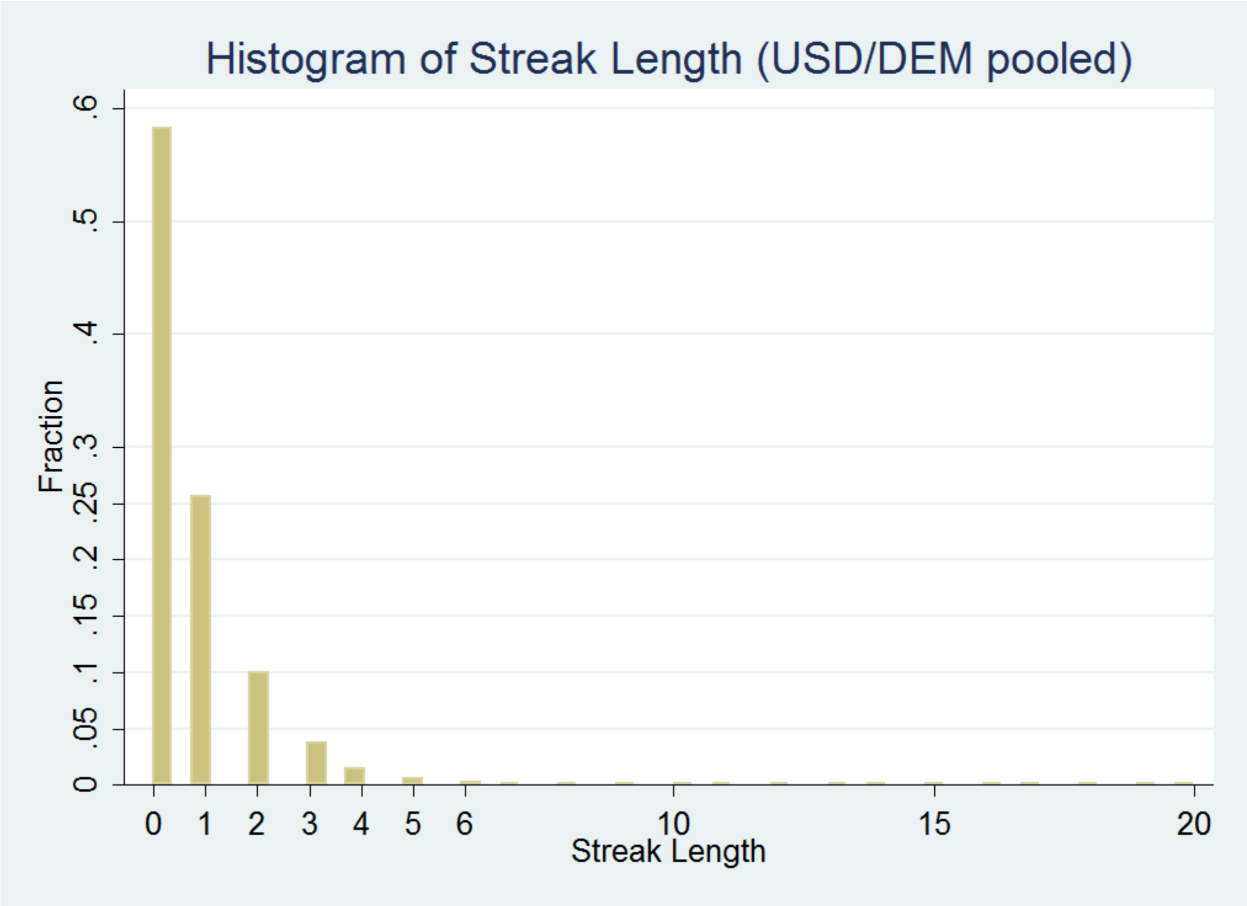
[Please insert Table A2 about here]

A.3. An Autoregressive Model

In this section, we report the results of AR(15) models mentioned in section 6, where x_t in Table A3 is $\Delta s_t$for column 1 and is $(s_{t+6} - s_{t+1})$ for column 2.

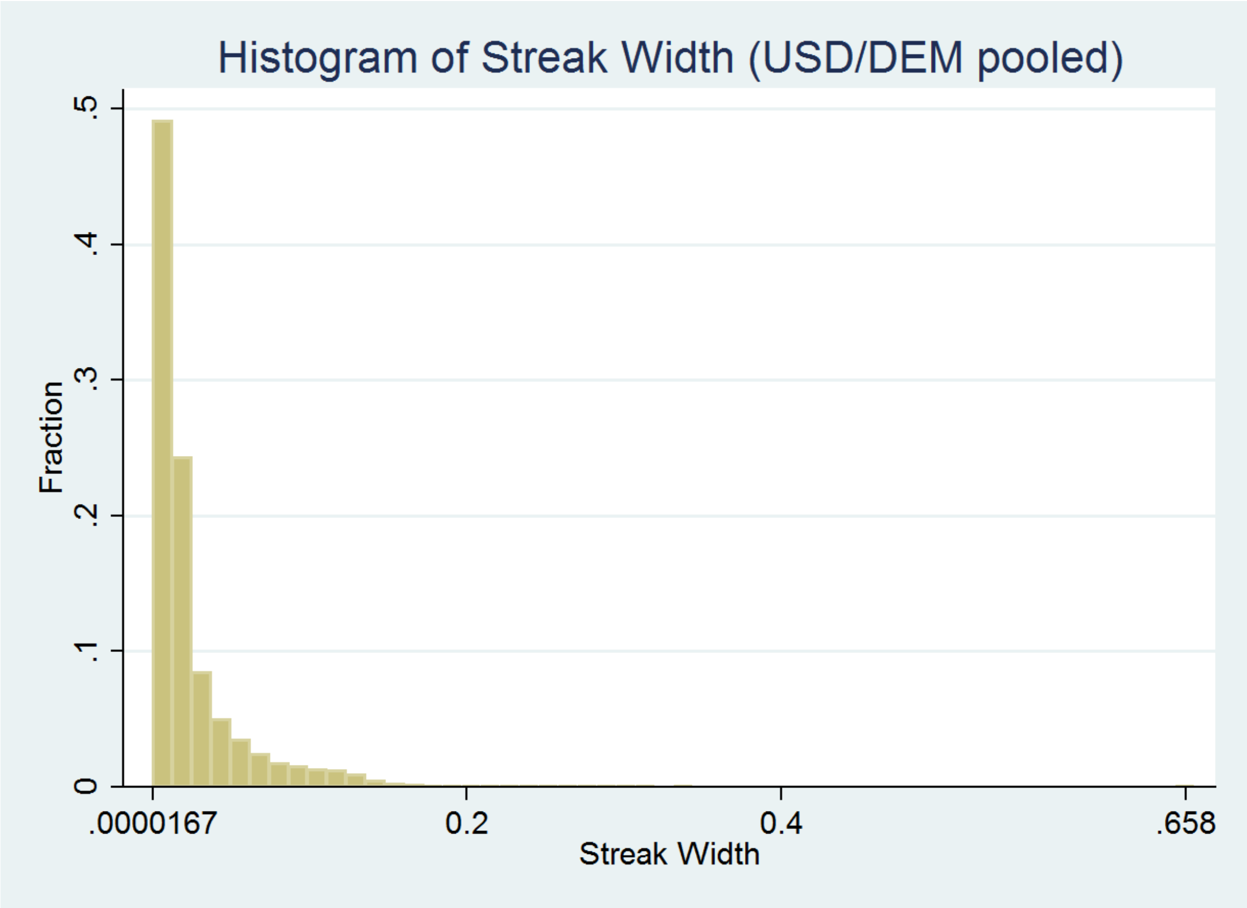
[Please insert Table A3 about here]

Figure 1: Histogram of Streak Length (USD/DEM pooled)



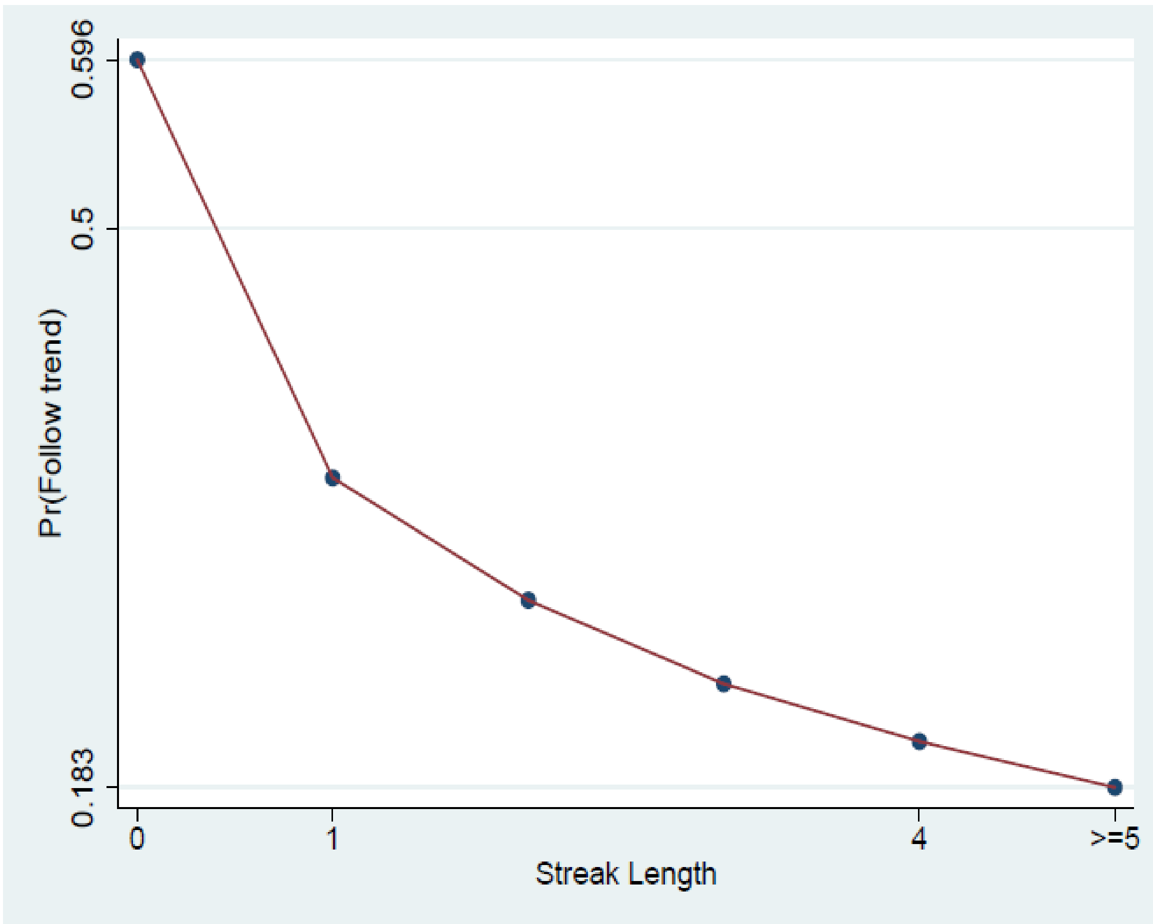
Source: The histogram is obtained from (7) using USD/DEM transaction data on the Reuters D2000-1 system.

Figure 2: Histogram of Streak Width (USD/DEM pooled)



Source: The histogram is obtained from (8) using USD/DEM transaction data on the Reuters D2000-1 system.

Figure 3: Predicted probability from equation (8) in Table 3



Source: The graph is obtained from (8) using USD/DEM transaction data on the Reuters D2000-1 system

Table 1: Data Format (USD/DEM)

Month	Day	Hour	Min	Sec	B/S	Ask	Bid
4	30	18	45	40	1	1.5326	-
4	30	18	46	23	0	-	1.5326
4	30	18	47	56	1	1.5328	-
4	30	18	48	22	1	1.533	-
4	30	18	49	53	1	1.5332	-
4	30	18	51	0	0	-	1.5327
4	30	18	52	34	1	1.5327	-
4	30	18	53	8	1	1.533	-
4	30	18	53	35	0	-	1.5329
4	30	18	54	21	1	1.5329	-
4	30	18	54	27	0	-	1.5333
4	30	18	55	10	0	-	1.533

This table shows the format of the data on the Reuters D2000-1 system. If B/S is equal to 1 there is a buy order and if B/S is equal to 0 there is a sell order.

Table 2: Summary Statistics

	Obs.	Mean	Std. Dev.	Min.	Max
Δs	256838	-1.12×10^{-7}	3.85×10^{-4}	-0.00663	0.00271
Streak length (k)	256838	0.700	1.0437	0	19
Streak width (w)	256838	0.0235	0.0292	2.98×10^{-6}	0.663

Source: USD/DEM transaction data from the Reuters D2000-1 system.

Table 3: Trend-following behaviour and streak length

	(7)	(7*)	(8)	(8*)
k_t	-0.343^{***}	-0.137^{***}		
	(-105.10)	(-105.10)		
Dk_1			-0.605^{***}	-0.233^{***}
			(-101.93)	(-108.05)
Dk_2			-0.800^{***}	-0.290^{***}
			(-91.42)	(-108.42)
Dk_3			-0.945^{***}	-0.324^{***}
			(-67.05)	(-89.97)
Dk_4			-1.054^{***}	-0.345^{***}
			(-45.23)	(-67.29)
Dk_{5+}			-1.148^{***}	-0.364^{***}
			(-37.15)	(-60.11)
Constant	0.171^{***}		0.242^{***}	
	(54.47)		(73.03)	
Pseudo-R^2	0.050		0.059	
Probability		0.473		0.474
Obs.	256838		256838	

Dependent variable: TF_t . Estimation method: probit. Columns numbered * are marginal effects. k = streak length. Dk_i=1 for k=i, =0 otherwise. Dk_{5+} =1 for k≥5, =0 otherwise. Z-statistics in parentheses: $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

Table 4: Trend-following Behaviour, Streak Length and Streak Width

	(9)	(9*)
Dk_1	-0.617***	-0.237***
	(-102.66)	(-109.06)
Dk_2	-0.816***	-0.300***
	(-92.36)	(-110.27)
Dk_3	-0.964***	-0.329***
	(-68.05)	(-92.38)
Dk_4	-1.075***	-0.350***
	(-46.04)	(-69.59)
Dk_{5+}	-1.171***	-0.368***
	(-37.82)	(-62.38)
$Streak\ width$	-1.204***	-0.479***
	(-13.68)	(-13.68)
Constant	0.277***	
	(66.43)	
Pseudo-R^2	0.0583	
Probability		0.474
Obs.	256838	

Dependent variable: TF_t . Estimation method: probit. Columns numbered * are marginal effects. k = streak length. Dk_i=1 for k=i, =0 otherwise. Dk_{5+} =1 for k≥5, =0 otherwise. Z-statistics in parentheses: $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

Table 5: Time Interval Effects

	(10a)	(10a*)	(10b)	(10b*)
Dk_1	-0.617***	-0.237***	-0.666***	-0.255***
	(-102.67)	(-109.08)	(-80.68)	(-86.37)
Dk_2	-0.816***	-0.295***	-0.863***	-0.309***
	(-92.35)	(-110.26)	(-71.10)	(-86.39)
Dk_3	-0.964***	-0.329***	-1.013***	-0.341***
	(-68.03)	(-92.35)	(-52.51)	(-73.29)
Dk_4	-1.076***	-0.350***	-1.107***	-0.357***
	(-46.03)	(-69.60)	(-35.20)	(-54.40)
Dk_{5+}	-1.166***	-0.367***	-1.224***	-0.378***
	(-37.35)	(-61.35)	(-29.39)	(-50.64)
w_t	-1.231***	-0.490***	-1.351***	-0.538***
	(-13.91)	(-13.91)	(-10.73)	(-10.73)
T_t	0.000041	0.000016	0.00061	0.00025
	(0.92)	(0.92)	(1.00)	(1.00)
$Dk_1 * T$	0.00024***	0.000094***		
	(2.76)	(2.76)		
$Dk_2 * T$	0.00025*	0.000098*		
	(1.93)	(1.93)		
$Dk_3 * T$	0.00028	0.00111		
	(1.34)	(1.34)		
$Dk_4 * T$	-0.00005	-0.000023		
	(-0.26)	(-0.26)		
$Dk_{5+} * T$	0.00107*	0.000425*		
	(1.66)	(1.66)		
TF_{t-1}				
Constant	0.276***		0.274***	
	(64.60)		(64.70)	
Pseudo-R^2	0.0584		0.0659	
Probability		0.474		0.475
Obs.	256838	256838	136118	136118

Dependent variable: TF_t . Estimation method: probit. Columns numbered * are marginal effects. k = streak length. Dk_i=1 for k=i, =0 otherwise. Dk_{5+} =1 for k≥5, =0 otherwise. w = streak width. Regression (12a) uses observations between the 25th and 75th percentiles of T only.
Z-statistics in parentheses: $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

Table 6: Microstructure Effects

	(11)	**(11*)**
$\boldsymbol{Dk_1}$	-0.670***	-0.256***
	(-107.14)	(-114.86)
$\boldsymbol{Dk_2}$	-0.833***	-0.300***
	(-94.78)	(-113.85)
$\boldsymbol{Dk_3}$	-0.968***	-0.329***
	(-68.86)	(-93.68)
$\boldsymbol{Dk_4}$	-1.070***	-0.349***
	(-46.23)	(-69.61)
$\boldsymbol{Dk_{5+}}$	-1.155***	-0.365***
	(-37.64)	(-61.27)
$\boldsymbol{w_t}$	-1.282***	-0.510***
	(-14.49)	(-14.49)
$\boldsymbol{TF_{t-1}}$	0.174***	0.0692***
	(32.86)	(32.93)
Constant	0.210***	
	(45.48)	
Pseudo-R^2	0.0613	
Probability		0.474
Obs.	256822	256822

Dependent variable: TF_t . Estimation method: probit. Columns numbered * are marginal effects. k = streak length. Dk_i=1 for k=i, =0 otherwise. Dk_{5+} =1 for k≥5, =0 otherwise. w = streak width.
Z-statistics in parentheses: $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

Table 7: Streaks and Exchange Rate Returns

Dependent variable:	**$1000\Delta s_{t+1}$**	**$1000\Delta s_{t+1}$**	**$100000(s_{t+6} - s_{t+1})$**
***k** sign**	-0.0198***	-0.0441***	0.0187
	(-28.25)	(-65.94)	(0.23)
w_t* sign		-4.555***	-0.484
		(-190.07)	(-0.17)
sign	-0.0869***	0.037***	-0.815***
	(-98.54)	(35.20)	(-6.51)
Constant	0.00015	-0.000257	-0.0599
	(0.20)	(-0.37)	(-0.73)
R^2	0.0713	0.186	0.0004
RMSE	0.37	0.35	41
Obs.	256822	256822	256742

The variable "sign" =1 if $\Delta s_t > 0$ and = -1 if $\Delta s_t < 0$.
Z-statistics in parentheses: $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

Table 8: Predictions (Part I)

Dependent variable:	**Model A**	**Marginal effect**	**Model B**	**Marginal effect**
Dk_1	-0.566***	-0.218***	-0.558***	-0.216
	(-90.46)	(-95.25)	（-23.85）	（-25.1）
Dk_2	-0.739***	-0.271***	-0.698***	-0.258
	(-80.43)	(-93.24)	（-15.15）	（-17.32）
Dk_3	-0.873***	-0.305***	-0.787***	-0.281
	(-60.14)	(-77.48)	（-11.35）	（-13.99）
Dk_4	-0.977***	-0.328***	-0.839***	-0.293
	(-41.22)	(-58.10)	（-8.98）	（-11.59）
Dk_{5+}	-1.0634***	-0.346***	-0.833***	-0.291
	(-33.90)	(-51.30)	（-6.28）	（-8.13）
w_t	1.532***	0.610***	4.887*	1.946
	(12.00)	(12.00)	（2.09）	（2.09）
Prediction	571.598***	227.565***	1337.569**	532.526
	(29.86)	(29.86)	（2.61）	（2.61）
Constant	0.251***		0.227***	
	(58.86)		（11.7）	
Pseudo-R^2	0.0612		0.058	
Probablity		0.474		0.475
Obs.	256838		256838	

Probit Models.

$$\textbf{Model A} = \frac{\textbf{Fitted values of the AR model about } \Delta s_{t+1}}{\textbf{sign}}$$

$$\textbf{Model B} = \frac{\textbf{Fitted values of the model}(\Delta s_{t+1} \textbf{ on sign}, \mathbf{k} * \textbf{sign}, \mathbf{w} * \textbf{sign})}{\textbf{sign}}$$

Table 9: Predictions (Part II)

Dependent variable:	**Model C**	**Marginal effect**	**Model D**	**Marginal effect**
Dk_1	-0.620***	-0.238***	-0.582***	-0.225
	（-103.07）	(-109.56)	（-86.4）	(-91.26)
Dk_2	-0.821***	-0.297***	-0.779***	-0.284
	（-92.84）	(-111.06)	（-82.88）	(-97.58)
Dk_3	-0.971***	-0.330***	-0.926***	-0.319
	（-68.48）	(-93.34)	（-63.64）	(-84.49)
Dk_4	-1.0817***	-0.352***	-1.0360***	-0.342
	（-46.31）	(-70.35)	（-43.91）	(-64.5)
Dk_{5+}	-1.181***	-0.370***	-1.131***	-0.361
	（-38.13）	(-63.41)	（-36.31）	(-57.96)
w_t	-1.252***	-0.499***	-0.710***	-0.283
	（-14.19）	(-14.19)	（-7.25）	(-7.25)
Prediction	-0.0611***	-0.0243***	0.0453***	0.0180
	（-16.99）	(-16.99)	（11.58）	(11.58)
Constant	0.276***		0.278***	
	（66.33）		（66.5）	
Pseudo-R^2	0.0591		0.0587	
Probablity		0.475		0.475
Obs.	256822		256838	

Probit model.

$$\textbf{Model C} = \frac{\textbf{Binomial Fitted values of the AR model about } \Delta s_{t+1}}{\textbf{sign}}$$

$$\textbf{Model D} = \frac{\textbf{Binomial Fitted values of the model}(\Delta s_{t+1} \textbf{ on sign}, \mathbf{k} * \textbf{sign}, \mathbf{w} * \textbf{sign})}{\textbf{sign}}$$

Table A1: Results for observations with infrequent trading only

	(A1*)	(A1)	(A2)	(A2*)	(A3)	(A3*)	(A4)	(A4*)
Dk_1	-0.20***	-0.53***	-0.58***	-0.22***	-0.54***	-0.21***	-0.60***	-0.23***
	(-27.46)	(-26.33)	(-42.87)	(-45.19)	(-26.55)	(-27.74)	(-43.37)	(-45.85)
Dk_2	-0.27***	-0.73***	-0.76***	-0.28***	-0.74***	-0.27***	-0.78***	-0.29***
	(-27.33)	(-23.86)	(-37.91)	(-44.16)	(-24.13)	(-27.82)	(-38.54)	(-45.27)
Dk_3	-0.30***	-0.84***	-0.88***	-0.31***	-0.86***	-0.31***	-0.91***	-0.32***
	(-21.16)	(-16.96)	(-27.00)	(-34.75)	(-17.26)	(-21.78)	(-27.61)	(-36.05)
Dk_4	-0.35***	-1.04***	-1.02***	-0.34***	-1.06***	-0.35***	-1.05***	-0.35***
	(-16.44)	(-11.36)	(-18.28)	(-26.24)	(-11.58)	(-17.03)	(-18.75)	(-27.46)
Dk_{5+}	-0.31***	-0.89***	-1.00***	-0.34***	-0.91***	-0.32***	-1.03***	-0.34***
	(-10.58)	(-8.05)	(-13.41)	(-19.16)	(-8.25)	(-11.03)	(-13.79)	(-20.14)
w_t					-0.91***	-0.36***	-1.25***	-0.45***
					(-3.73)	(-3.73)	(-7.26)	(-7.26)
Constant		0.235	0.241		0.272		0.286	
		(21.04)	(32.11)		(18.40)		(29.41)	
Pseudo-R^2		0.0450	0.0517		0.0454		0.0525	
Probability	0.491			0.484		0.491		0.484
Obs.		21780	48933		21780		48933	

* marginal effects.

Table A2: Microstructure Effects

	(11’)	(11’*)
Dk_1	-0.667***	-0.254***
	(-106.54)	(-114.19)
Dk_2	-0.835***	-0.301***
	(-94.258)	(-113.31)
Dk_3	-0.975***	-0.331***
	(-69.18)	(-94.48)
Dk_4	-1.073***	-0.349***
	(-46.41)	(-70.06)
Dk_{5+}	-1.153***	-0.365***
	(-37.66)	(-61.23)
w_t	-1.288***	-0.513***
	(-14.55)	(-14.55)
TF_{t-1}	0.173	0.0688
	(32.79)	(32.86)
TF_{t-2}	0.0173	0.00691
	(3.35)	(3.35)
TF_{t-3}	0.0213	0.00846
	(4.15)	(4.15)
TF_{t-4}	0.00716	0.00285
	(1.4)	(1.4)
TF_{t-5}	0.00654	0.00260
	(1.29)	(1.29)
TF_{t-6}	0.00208	0.000829
	(0.41)	(0.41)
TF_{t-7}	0.00790	0.00314
	(1.55)	(1.55)
TF_{t-8}	0.00435	0.00173
	(0.86)	(0.86)
TF_{t-9}	0.00793	0.00316
	(1.56)	(1.56)
TF_{t-10}	0.000966	0.000384
	(0.19)	(0.19)

Constant	0.174***	
	(20.76)	
Pseudo-R^2	0.0613	
Probability		0.474
Obs.	256822	256822

Dependent variable: TF_t . Estimation method: probit. Columns numbered * are marginal effects. k = streak length. Dk_i=1 for k=i, =0 otherwise. Dk_{5+} =1 for k≥5, =0 otherwise. w = streak width.

The coefficients of the lags of TF are not reported.
Z-statistics in parentheses: $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

Table A3: An autoregressive model of exchange rate returns

AR(15) model	x_t is Δs_t	x_t is $(s_{t+6} - s_{t+1})$
x_{t-1}	-0.759	0.184
	(-384.58)	(95.33)
x_{t-2}	-0.594	0.126
	(-239.45)	(64.35)
x_{t-3}	-0.464	0.102
	(-169.14)	(51.62)
x_{t-4}	-0.365	0.0783
	(-126.48)	(39.49)
x_{t-5}	-0.287	-0.684
	(-96.5)	(-345.52)
x_{t-6}	-0.228	0.161
	(-75.23)	(71.8)
x_{t-7}	-0.177	0.120
	(-57.92)	(53.01)
x_{t-8}	-0.136	0.101
	(-44.31)	(44.75)
x_{t-9}	-0.106	0.0784
	(-34.58)	(34.69)
x_{t-10}	-0.0790	-0.429
	(-26.13)	(-191.57)
x_{t-11}	-0.0604	0.0899
	(-20.29)	(45.4)
x_{t-12}	-0.0446	0.0709
	(-15.43)	(35.8)
x_{t-13}	-0.0288	0.0670
	(-10.52)	(33.9)
x_{t-14}	-0.0163	0.0531
	(-6.56)	(27.04)
x_{t-15}	-0.00948	-0.203
	(-4.8)	(-104.72)
Constant	-5.16e-07	-6.60e-07
	(-0.85)	(-0.99)
$\mathbf{R^2}$	0.37	0.34
Obs.	256598	256502